DiffusionDB: A Large-scale Prompt Gallery Dataset for Text-to-Image Generative Models

Year: 2022 | *Citations: 377* | *Authors: Zijie J. Wang, Evan Montoya, David Munechika*

Abstract

With recent advancements in diffusion models, users can generate high-quality images by writing text prompts in natural language. However, generating images with desired details requires proper prompts, and it is often unclear how a model reacts to different prompts or what the best prompts are. To help researchers tackle these critical challenges, we introduce DiffusionDB, the first large-scale text-to-image prompt dataset totaling 6.5TB, containing 14 million images generated by Stable Diffusion, 1.8 million unique prompts, and hyperparameters specified by real users. We analyze the syntactic and semantic characteristics of prompts. We pinpoint specific hyperparameter values and prompt styles that can lead to model errors and present evidence of potentially harmful model usage, such as the generation of misinformation. The unprecedented scale and diversity of this human-actuated dataset provide exciting research opportunities in understanding the interplay between prompts and generative models, detecting deepfakes, and designing human-AI interaction tools to help users more easily use these models. DiffusionDB is publicly available at: https://poloclub.github.io/diffusiondb.